Objectives

- Describe methods of capturing, selecting, managing and exchanging data

- Describe what is meant by transaction processing and ACID (Atomicity, Consistency, Isolation, Durability)

- Describe what is meant by record locking and why it is necessary in a multi-user database
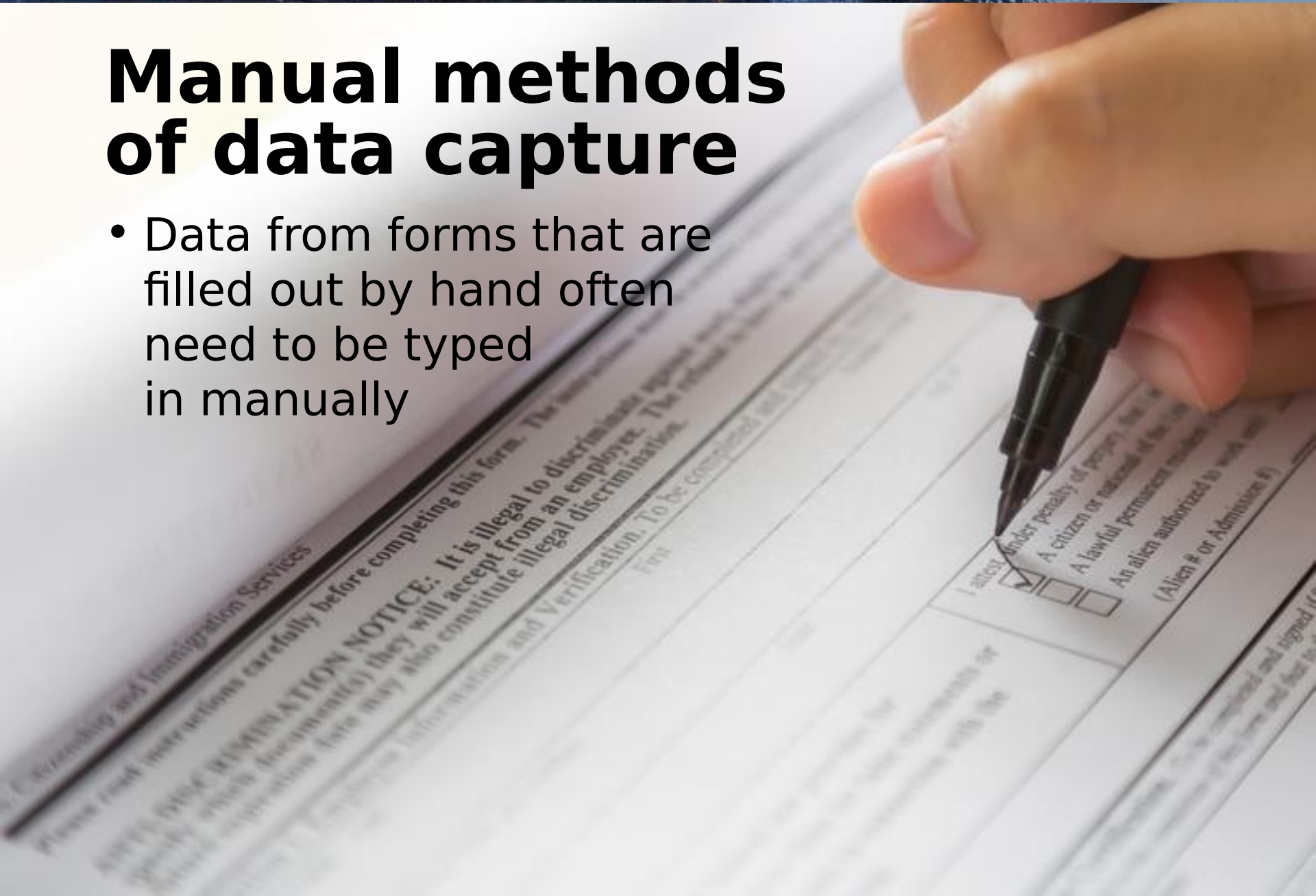
- Describe what is meant by redundancy

Capturing data

- How many methods of data capture can you name?

 - Suggest an application for each of them

Manual methods of data capture

- Data from forms that are filled out by hand often need to be typed in manually

Automated methods

- These include

  - smart card readers

  - barcode readers

  - scanners

  - optical character recognition (OCR)

  - optical mark recognition (OMR)

  - magnetic ink character recognition

  - sensors

- Suggest a use for each of these methods
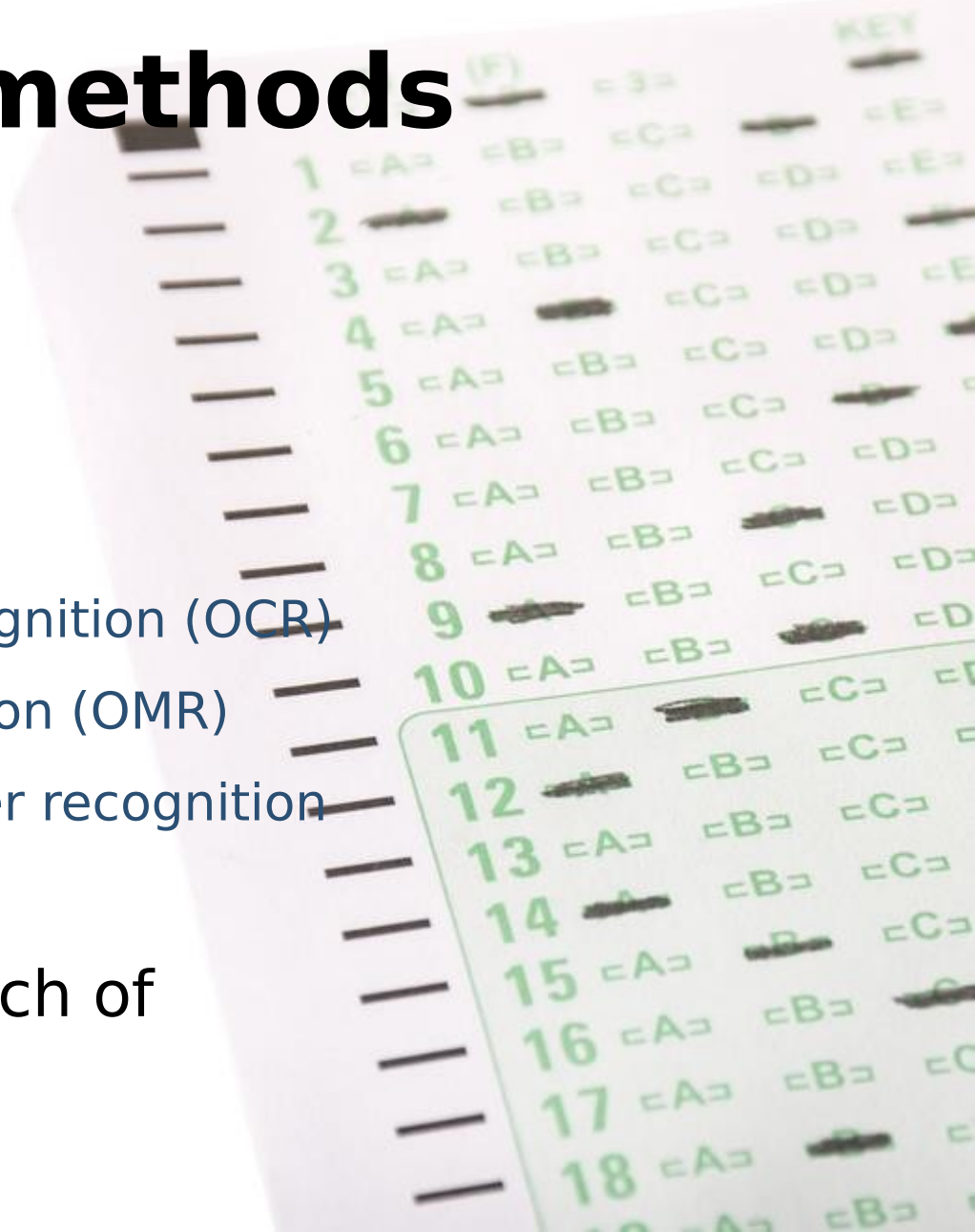
Inputting data

- Once data has been collected, it can be transferred to a database

  - automatically, using the DBMS software

  - by typing it in using a customised form

  - importing it from a spreadsheet or file

  - using EDI (Electronic Data Interchange) – this is used to transfer data between one computer system and another

PG ONLINE

EDI

- Electronic Data interchange (EDI) is the computer-to-computer exchange of documents such as purchase orders, invoices and shipping documents between two companies or business partners

- It replaces post, email or fax

- All documents must be in a standard format so that the computer can understand them

- EDI translation software may be used to translate the EDI format so the data can be input directly to a company database

PG ONLINE

Selecting data

- Using SQL or, for example, Query By Example in Access, data satisfying specific criteria can be selected and ordered

  - It can then be used in reports, letters, mailing labels, to print out barcodes, etc.

Transaction processing

- In the context of databases, a single logical operation is defined as a transaction

- It may consist of several operations; for example, a customer order may consist of several order lines…

  - all of which must be processed…

  - the quantity of each product adjusted on the stock file…

  - credit card details checked…

  - payment accepted or rejected

- What happens if the stock file has been updated and the system crashes before the payment is processed?

Worksheet 6

- Do the questions in **Task 1**

ACID

- ACID stands for Atomicity, Consistency, Isolation, Durability

- This is a set of properties to ensure that the integrity of the database is maintained under all circumstances

- It guarantees that transactions are processed reliably

Atomicity

- This property requires that a transaction is processed in its entirety or not at all

- In any situation, including power cuts or hard disk crashes, it is not possible to process only part of a transaction

PG ONLINE

Consistency

- This property ensures that no transaction can violate any of the defined validation rules

- Referential integrity, specified when the database is set up, will always be upheld

PG ONLINE

Isolation

- The Isolation property ensures that concurrent execution of transactions leads to the same result as if transactions were processed one after the other

- This is crucial in a multi-user database

Durability

- This ensures that once a transaction has been committed, it will remain so, even in the event of a power cut

- As each part of a transaction is completed, it is held in a buffer on disk until all elements of the transaction are completed

- Only then will the changes to the database tables be made

Potential problems with multi-user databases

- What are the potential problems?

- Suppose several people are simultaneously trying to reserve the last few seats on an aeroplane … what could happen?

PG ONLINE

Multi-user databases

- Allowing multiple users to simultaneously access a database could potentially cause one of the updates to be lost

- For example, consider this procedure:

  - When an item is to be updated, the entire block in which the record is located is read into the user's own local memory at the workstation

  - When the record is saved, the block is rewritten to the file server

- What can go wrong?
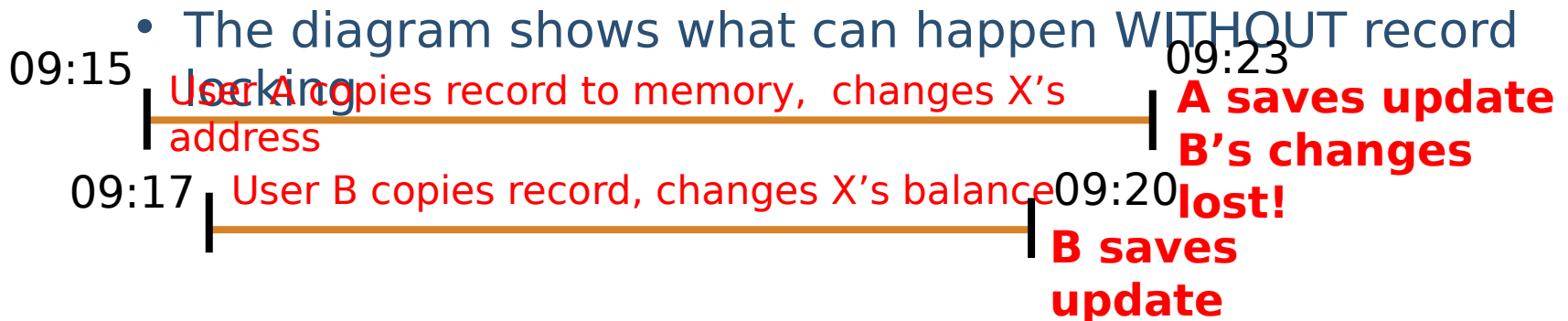
PG ONLINE

Imagine the scenario…

- User A accesses a customer record and it is copied to their local storage

- User A starts to alter the customer's address

Imagine the scenario...

- User B accesses the same customer record, alters the credit limit and saves the record

- User A completes the change of address, and saves the record

- What is the state of the record?

Record locking

- Record locking prevents simultaneous access to objects in a database in order to prevent updates being lost or inconsistencies in the data arising

- The diagram shows what can happen WITHOUT record locking

09:15   User A copies record to memory, changes X's address   09:23 **A saves update**

09:17   User B copies record, changes X's balance 09:20 **B's changes lost!**

B saves update

- Using record locking, a record is locked when a user retrieves it for editing or updating

- Anyone else attempting to retrieve it is denied access until the transaction is completed or cancelled

PG ONLINE

Problems with record locking

- If two users are attempting to update two records, a situation can arise in which neither can proceed, known as deadlock

  - Ken is attempting to make a transfer from Customer A's account to Customer B's account

  - Meanwhile, Paula is attempting to make a transfer from Customer B's account to Customer A's account

PG ONLINE

Deadlock!

- Ken locks Customer A's record

- Tries to access Customer B's record

- Waits…

- Paula locks Customer B's record

- Tries to access Customer A's record

- Waits…

PG ONLINE

Serialisation

- The Database Management System (DBMS) must prevent such situations from arising

- Serialisation ensures that transactions do not overlap in time and therefore cannot interfere with each other or lead to updates being lost

- Serialisation techniques include:

  - Timestamp ordering

  - Commitment ordering

PG ONLINE

Timestamp ordering

- Every object in the database has a read timestamp and a write timestamp

- These are updated whenever an object is read
or written

Timestamp ordering

- When a user tries to save an update, if the read timestamp is not the same as it was when they started the transaction, the DBMS knows another user has accessed the same object
  - The transaction will be cancelled and a message "*Update unsuccessful*" sent to the user

Object timestamp

0905

User A copies record to memory, changes X's balance

Object timestamp

0910… A cannot save update

Object timestamp

0906

User B copies record, changes X's balance

Object timestamp

0910… B saves update

PG ONLINE
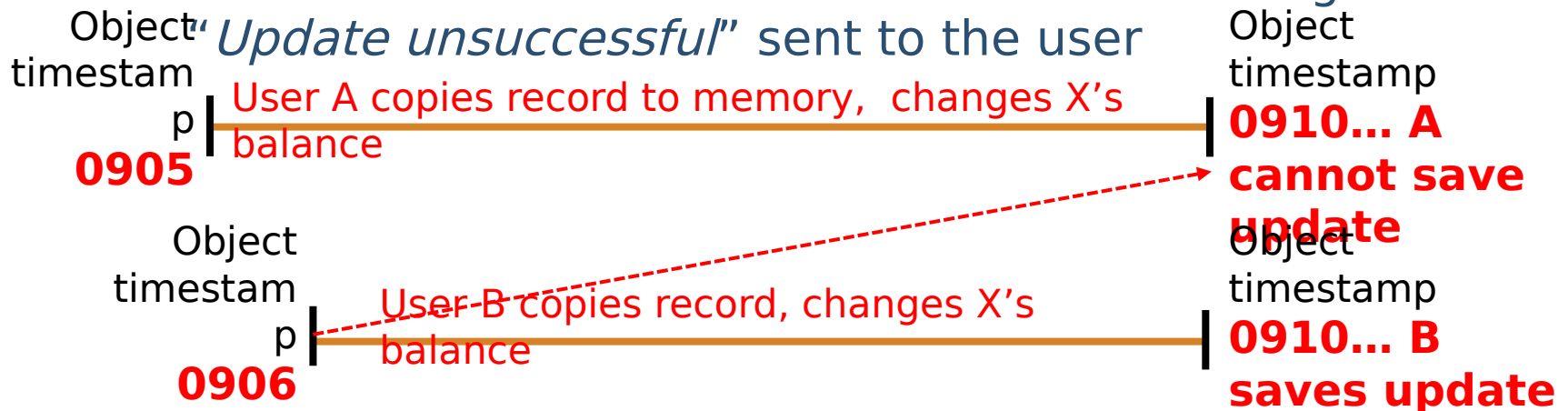
Commitment ordering

- This is another serialisation technique to ensure that no transactions are lost if two clients are simultaneously trying to update a record

- Transactions are ordered in terms of their dependencies on one another as well as the time they were initiated

  - It can be used to prevent deadlock by blocking one request until another is completed

PG ONLINE

Redundancy

- Many organisations cannot afford to have their computer systems go down for even a short time

  - Imagine the chaos if the air traffic control system goes down…

Redundancy

- Many organisations have built-in redundancy in their computer systems

- Duplicate hardware, located in different geographical areas, mirrors every transaction that takes place on the main system

- If this fails, the backup system automatically takes over

Worksheet 6

- Do the questions in **Task 2** on the worksheet

Plenary

- You should be able to describe:

  - methods of capturing, selecting, managing and exchanging data

  - what is meant by transaction processing and ACID (Atomicity, Consistency, Isolation, Durability)

  - what is meant by record locking and why it is necessary in a multi-user database

  - what is meant by redundancy

Copyright

PG ONLINE